PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (Error! Hyperlink reference not valid.) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The tuberculosis care cascade in Zambia - identifying the gaps in order to improve outcomes: a population-based analysis
AUTHORS	Lungu, Patrick; Kerkhoff, Andrew ; Kasapo, Clara; Mzyece, Judith; Nyimbili, Sulani; Chimzizi, Rhehab; Silumesi, Andrew; Kagujje, Mary; Subbaraman, Ramnath; Muyoyeta, Monde; Malama, Kennedy

VERSION 1 – REVIEW

REVIEWER	Loutet, Miranda Public Health England
REVIEW RETURNED	12-Nov-2020

GENERAL COMMENTS	Really interesting paper and application of these novel technique of evaluating the cascade of care for people with TB using minimal surveillance data. It's great to see how techniques from papers in South Africa and India can be applied to other settings. However, there are some fundamental problems with the datasets and estimation methods as they are currently presented: - Firstly, it's hard to follow exactly how each step was calculated and what dataset was used. I suggest clarifying this in two ways. First, add which data sets were used at each step in all tables (main text of paper and supplementary tables). Secondly, organize the methods section so that it is clear exactly how each step is calculated and which dataset is used (following more closely how they did in Subbaraman et al., 2016 and Naidoo et al., 2017). - Consider what the appropriate denominator is throughout all analyses and report that clearly, instead of saying "overall proportion" and "relative proportion". For example, the appropriate denominator for reporting drug sensitive or drug resistant cases is out of those with any drug susceptibility testing. Also, present 95% confidence intervals for proportions within the text of the paper so that rages are very clear to readers. - Be really clear about all assumptions and caveats upfront – do not wait until the limitations in the discussion section to bring these up. The biggest caveat is that you are assuming perfect data collection. So the 40% of cases who did not seek care or undergo microbiologic TB testing, how can you be sure this is not due to missing data? As you move along the cascade and the numbers drop, again you cannot be sure these are the same people you are following throughout, they could be completely different people being captured in the system at different times. I understand this is an assumption you are making, but you need to be very clear

	about this because how it is written right now it sounds like you have a certain number of cases and you follow them individually through their care cascade and know when they are dropping out. - Following from the last point about assumptions and caveats, can you say that your results are generalizable? Are you worried that you are missing cases because you only use data from 4 of 10 provinces (line 128-129). Also, why is this province data for 2017 (line 130), when the aim of the paper is to present the gaps in overall care cascade in 2018. I would suggest reconsidering if the authors can state that this analysis relates to 2018 because the only number that is relevant to 2018 is the WHO estimates. Another issue with stating that the analysis refers to the 2018 cascade of care is that the data to inform the drug resistant rates is from 2008, so does not align to more current drug resistant rates (which have been greatly increasing). - How is drug-sensitive (DS) defined? Are people with isoniazid monoresistance included in the DS group? Why was MDR-TB not considered in the analysis? Although this type of analysis is generally considered exempt from full ethics board review, it still should have an IRB review stating such. There is no need for the Patients and public involvement section.

REVIEWER	Chikovore, Jeremiah Human Sciences Research Council
REVIEW RETURNED	23-Nov-2020

GENERAL COMMENTS	1. The authors' conclusion highlights the need to strengthen health systems and implement active care-seeking strategies. This is certainly crucial. The authors could also additionally frame their recommendations in the context of current debates whereby health status, access, and behaviors are influenced by social as well as structural factors, which often tend to work in combination. As currently framed, the emphasis seems to be on health services and how they are provided to individuals and communities. I am also curious regarding the significance of gender to the cascade. If the authors do not feel it is necessary to mention, it is their call based on their data or their priorities for this analysis; however, gender/sex is a variable that is usually included and would be crucial to mention given its salience to TB epidemiology. 2. The authors also indicate they intended to identify the largest gaps. This is certainly important to pay attention to areas with the greatest gaps. I would also consider/propose that even stages where the cascade drops seem 'small' may still be very crucial to TB control, therefore that all fall-out from the cascade should be a priority. However, depending on what the authors intend to emphasize, their pitch of the value of cascade analyses (and of this specific analysis) is certainly their call to make. 3. The authors might also want to look at qualitative research studies carried out in high burden settings to get insights into possible explanations for their results and observations. 4. The authors also note the need for collaborative TB/HIV activities in their conclusion. They refer to a WHO reference from

	2012. In light of developments in the HIV field in more recent years, the authors could consider acknowledging this through a passing reference to these changes and their potential implications to their findings 5. Line 403 estimate [54]. “Due to a lack of a unique national patient identifier, we were unable to link individual patient outcomes as they progressed through the TB care cascade; where possible”. Are the authors suggesting if they had personal IDs for the patients, they would have tracked patients as they moved through the cascade? 6. Line 4 – I suggest removing ‘among individuals’ 7. Line 46 – add ‘coma’ between ‘Zambia’ and ‘have’ 8. Line 47 – suggest citing the latest Global TB report, if possible. 9. Line 53-55: Review grammar for sentence construction and flow. 10. Line 60-62: Review the construction of the sentence 11. Line 66 – Consider not making rifampicin-susceptibility a single word; same for HIV-status 12. Line 73 – Consider removing ‘Living in its Provinces’ 13. Line 7 - There is a word missing 14. Line 81 - Indicate the basis of your estimation that a small number of patients are detected and managed in the private sector; 15. Line 88 – Consider changing ‘multi-drug resistant’ to become ‘multidrug-resistant’ 16. Could the authors elaborate why information from only 4 provinces was used to draw national programmatic data? 17. Line 133: To improve intelligibility, please consider inserting coma just before ‘and smear’ 18. Line 184 - Consider elaborating “Xpert tests sent each year” - to where? 19. Line 219: Sentence starting ‘The majority…’ is missing a word. 20. Line 227-229 – Consider skipping part after ‘HIV,” 21. Line 257: If ‘treatment outcomes’ is referring to ‘Treatment completion rates’, perhaps authors could consider phrasing consistently 22. Line 297 – ‘This has been…”; please specify ‘what’ has been 23. Line 299 - check gramma 24. Line 350 – Consider correcting gramma 25. Line 355: Consider rephrasing to, especially regarding this part: “presenting to and in-care” 26. Please define LFTU at first use 27. The outcomes figure shows overall improvement, but a downward trend for 3 of the 5 measures. Is this something the authors may consider commenting on even in passing; if there is something to read into it?

VERSION 1 – AUTHOR RESPONSE

REVIEWER #1

1.Really interesting paper and application of these novel technique of evaluating the cascade of care for people with TB using minimal surveillance data. It’s great to see how techniques from papers in South

Africa and India can be applied to other settings. However, there are some fundamental problems with the datasets and estimation methods as they are currently presented:

Response: We thank the reviewer for the positive feedback and their thoughtful review. We have endeavored to address all concerns raised and by the reviewer and to incorporate their suggestions into the revised manuscript.

2.Firstly, it's hard to follow exactly how each step was calculated and what dataset was used. I suggest clarifying this in two ways. First, add which data sets were used at each step in all tables (main text of paper and supplementary tables). Secondly, organize the methods section so that it is clear exactly how each step is calculated and which dataset is used (following more closely how they did in Subbaraman et al., 2016 and Naidoo et al., 2017).

Response: We thank the reviewer for this comment and agree that the methods section can be further revised to improve clarity. In an effort to enhance understanding of the approach to estimates and the data sources informing estimates, we have added a new Table 1 that closely mirrors the approach used by Subbaraman et, 2016. We hope that this table provides further clarity. Because many of the steps required multiple calculations potentially informed by multiple data sources, it would be very challenging to clearly note which (and how) data sources informed the estimates in the Tables 2 and 3 (original Tables 1 and 2). We ultimately felt that this was the clearest way to convey the information requested by the reviewer and compliment the very detailed analysis summarized in the Supplementary Appendix. Additionally, in our Supplementary Appendix, which closely mirrors the approach/presentation utilized by Naidoo et all, 2017, we have added further information on data sources utilized whenever applicable. We have also made multiple revisions throughout the methods section taking into account the reviewer's point in an effort to further clarify the approach and data sources utilized.

3.Consider what the appropriate denominator is throughout all analyses and report that clearly, instead of saying "overall proportion" and "relative proportion". For example, the appropriate denominator for reporting drug sensitive or drug resistant cases is out of those with any drug susceptibility testing. Also, present 95% confidence intervals for proportions within the text of the paper so that rages are very clear to readers.

Response: We thank the reviewer this comment. We believe that presenting proportions in both ways has merit. The proportion relative to the total number of TB cases provides insight as to how many patients progress to each step of the cascade, while the proportion relative to the prior step may have relevance for specific quality improvement efforts. For the purposes of reporting the results, we have removed mention of "overall proportion and relative proportion" and only described proportions relative the total TB burden as this mirrors the presentation utilized by Subbaraman et al., 2016 and Naidoo et al., 2017 and improves comparability to their results; however, because we believe that both proportions utilizing different denominators provide complementary insight, we have retained both in Table 2 (previously Table 1). With reference to presenting 95%CI confidence intervals for proportions, we agree that this would be useful. However, uncertainty around case estimates at each step represent both ranges as well as 95% confidence intervals – we are not aware of a statistically valid way to calculate 95% confidence intervals for proportions using such estimates and are thus not able to present them. In an effort to make the uncertainty around estimates as clear as possible, we have noted the range, 95% confidence value (or specified the value as an exact value) where applicable throughout the revised results section.

4.Be really clear about all assumptions and caveats upfront – do not wait until the limitations in the discussion section to bring these up. The biggest caveat is that you are assuming perfect data collection. So the 40% of cases who did not seek care or undergo microbiologic TB testing, how can you be sure this is not due to missing data? As you move along the cascade and the numbers drop, again you cannot be sure these are the same people you are following throughout, they could be completely different people being captured in the system at different times. I understand this is an assumption you are making, but you need to be very clear about this because how it is written right now it sounds like you have a certain number of cases and you follow them individually through their care cascade and know when they are dropping out.

Response: We thank the reviewer for this point and we agree that it is important to be up front and transparent about assumptions/caveats. We acknowledge that the "denominator-denominator" linked approach (Haber, et al. Curr Opin HIV AIDS 2016;11(1):102-8.) in which the same individuals are followed through each step of the cascade would be less subject to bias, but such data was not available. Nonetheless, when such data are not available, we believe that using routine programmatic data, as we have done in this manuscript, allows for a reasonable understanding of the relative importance and scale of gaps in the TB care cascade. We have highlighted these two key points in the revised methods section, *"It should be noted that several steps of the cascade utilized exact numbers from aggregated facility-level programmatic data (steps 3, 4, and 5); forthe purposes of these analyses, data were assumed to be accurate and complete; however, such data may be incompletely recorded and a small proportion may be entered incorrectly - estimates of uncertainty around exact values from programmatic data were unavailable. Furthermore, unique patient identifiers are not available within Zambia's NTP and thus this analysis does not present a cohort of individuals that were tracked through each step of the TB care cascade; while we assumed for the purposes of this analysis that the same patients were being characterized at each step of the cascade, one cannot exclude the possibility that different individuals are being captured at different steps of the care cascade."*

5.Following from the last point about assumptions and caveats, can you say that your results are generalizable? Are you worried that you are missing cases because you only use data from 4 of 10 provinces (line 128-129). Also, why is this province data for 2017 (line 130), when the aim of the paper is to present the gaps in overall care cascade in 2018. I would suggest reconsidering if the authors can state that this analysis relates to 2018 because the only number that is relevant to 2018 is the WHO estimates. Another issue with stating that the analysis refers to the 2018 cascade of care is that the data to inform the drug resistant rates is from 2008, so does not align to more current drug resistant rates (which have been greatly increasing).

Response: We thank the reviewer for highlighting these important points. We do believe that our results are generalizable to the National TB Program in Zambia for 2018. We also believe that it is fair to describe this as a care cascade for the year 2018 because the core incidence estimates as well as diagnosis, notification and treatment numbers are all from 2018 programmatic data. The 2017 data from 4 of 10 provinces informed estimates regarding the <u>proportion of patients lost-to-follow-up</u>; thus, there is not a risk of 'missing cases.' In response to the reviewer's point as well as that made by reviewer #2, we have added further details to the manuscript regarding this dataset, *"This helped to further refine estimates for Steps 2 and 3 by accounting for and removing duplicate patients (Supplementary Appendix). Patient-level data was only available from 4 provinces of 10 provinces; however, they account for nearly 60% of Zambia's national TB notifications and the range of socioeconomic characteristics of individuals as well as their access to healthcare services*

are representative of the other 6 provinces. Unfortunately, robust data from 2018 to inform these estimates were unavailable – thus, we utilized 2017 data because it was well-characterized and temporally close to the year for which we sought to characterize the TB care cascade."

We understand and appreciate the reviewer's point about the year in which the most recent national survey on TB drug resistance in Zambia was conducted. We intentionally utilized this data in order to ground estimates of rifampicin-resistant TB in empiric data, rather using WHO incidence estimates of rifampicin-resistance, which may in part be informed by expert onion. However, in recognition of increased TB drug resistance rates, we intentionally chose higher end estimates from the most recent survey using MTBDRplus results – 2.4% (95%CI 1.2-3.6) to more closely align with WHO incidence estimates for 2018 (2.8%; range 2.5-3.1). We have noted this in the revised methods section, "***this source was chosen in order to ground estimates of RR-TB in empiric data, however, higher-end estimates from the latest Zambian national survey of TB drug resistance in 2008 were used to more closely align with WHO incidence estimates for RR-TB in 2018."***

Furthermore, to note the limitation that some data sources for the TB care cascade were not from 2018, we have mentioned this in the revised limitations section, *"Additionally, because core incidence, diagnosis, notification and treatment numbers are from 2018, we feel our analysis accurately represents the national TB care cascade in 2018; however, PTLTFU estimates were informed by patient-level data from 2017 and the proportion of cases with rifampicin resistance were informed by higher-end estimates from the most recent national drug resistance survey conducted in 2008 [20]. An updated drug resistance survey is currently underway and will provide new estimates that will better guide programmatic priorities."*

6.How is drug-sensitive (DS) defined? Are people with isoniazid monoresistance included in the DS group? Why was MDR-TB not considered in the analysis?

Response: We thank the reviewer for seeking clarification and in response, we have added greater clarity to the methods section. It is possible that INH-mono-resistant TB was included in the DS-TB group as TB drug susceptibility testing is not routinely performed unless rifampicin resistance is identified. This reflects the reality of many TB programs in high burden settings and concords with WHO recommendations. MDR-TB was indeed considered in the analysis, but was included under the umbrella definition of RR-TB of which the majority are likely MDR-TB cases as opposed to RIF-monoresistance. To further clarify these points, the methods have been revised as follows, "*DS-TB was defined as any TB case without known rifampicin resistance; thus, there is a possibility that patients with other forms of drug-resistance, including isoniazid monoresistance may have been included in this definition. However, unless rifampicin resistance is detected, TB drug susceptibility testing is not routinely performed in Zambia – this reflects the clinical reality of many high burden TB settings."*

7. Although this type of analysis is generally considered exempt from full ethics board review, it still should have an IRB review stating such.

Response: We appreciate the reviewer's comment. We have revised the wording to reflect that the study analysis was reviewed and received an exempt status, "*Because this was a retrospective, population-level analysis without the use of any patient identifiers, the University of Zambia Biomedical Research Ethics Committee determined that this analysis met the criteria for exempt-status."*

8. There is no need for the Patients and public involvement section.

Response: Thank you for this suggestion. However, we believe it is an editorial requirement to include this section and an appropriate statement. We defer to the editor as whether this section can/should remain.

REVIEWER #2

9. The authors' conclusion highlights the need to strengthen health systems and implement active care-seeking strategies. This is certainly crucial. The authors could also additionally frame their recommendations in the context of current debates whereby health status, access, and behaviors are influenced by social as well as structural factors, which often tend to work in combination. As currently framed, the emphasis seems to be on health services and how they are provided to individuals and communities. I am also curious regarding the significance of gender to the cascade. If the authors do not feel it is necessary to mention, it is their call based on their data or their priorities for this analysis; however, gender/sex is a variable that is usually included and would be crucial to mention given its salience to TB epidemiology.

Response: First, we would like to thank the reviewer for their time and thoughtful comments. Second, we agree that a greater discussion of patient's health-seeking and care engagement behaviors, including barriers to engagement and retention would further enrich and strengthen the discussion. We have gone through the discussion with a focus on this point and revised it accordingly. With regard to the reviewer's point on characterizing the cascade according to sex, we agree that this would be extremely informative data. Unfortunately, sex-disaggregated data sources were not available that would have allowed for us to estimate each step of the cascade for men and women. We have noted this in our revised limitations section along with the importance of applying a gender-lens to TB data collection as well as relevant gender-stratified data from Zambia's most recent National TB prevalence survey. "*Given the potential importance of gender to TB epidemiology and potential differential health-seeking behaviors and access to TB services, we sought to characterize the TB care cascade among men and women. For example, the prevalence of TB among men in Zambia's first national TB prevalence survey in 2013/2014 was almost twice as high as that among women (833 vs. 487 cases per 100,000 persons) and men with presumptive TB were likely to have sought care for their symptoms than women (31.4% vs. 38.4%). Unfortunately, sex-disaggregated data sources were not available that would have allowed for each step of the cascade to be estimated. It is important that TB programs collect sex-disaggregated diagnostic and treatment data to help ensure equity in access and treatment benefits."*

10.The authors also indicate they intended to identify the largest gaps. This is certainly important to pay attention to areas with the greatest gaps. I would also consider/propose that even stages where the cascade drops seem 'small' may still be very crucial to TB control, therefore that all fall-out from the cascade should be a priority. However, depending on what the authors intend to emphasize, their pitch of the value of cascade analyses (and of this specific analysis) is certainly their call to make.

Response: We thank the reviewer for this comment and agree with their point. To reflect this, we have slightly modified the background within the abstract section, *"We undertook a care cascade analysis to quantify gaps in care and align TB program improvement measures with areas of need."*

11.The authors might also want to look at qualitative research studies carried out in high burden settings to get insights into possible explanations for their results and observations.

Response: We thank the reviewer for this suggestion. We have added several qualitative references to the revised manuscript that we believe has further enriched the discussion section.

12.The authors also note the need for collaborative TB/HIV activities in their conclusion. They refer to a WHO reference from 2012. In light of developments in the HIV field in more recent years, the authors could consider acknowledging this through a passing reference to these changes and their potential implications to their findings

Response: We thank the reviewer for this suggestion. While this is an older reference, it is the most recent, applicable WHO guidance on this topic. In the discussion section we discuss many diagnostic developments that are applicable to PLHIV, including urine LAM and Xpert Ultra. If there are specific developments within the HIV field that the reviewer feels is explicitly missing and could further strengthen the manuscript through incorporation, we would be very happy to include this information.

13.Line 403 estimate [54]. “Due to a lack of a unique national patient identifier, we were unable to link individual patient outcomes as they progressed through the TB care cascade; where possible”. Are the authors suggesting if they had personal IDs for the patients, they would have tracked patients as they moved through the cascade?

Response: We thank the reviewer for seeking clarification of this point. We were trying to make clear that this was not a closed cohort in which individuals were tracked through each step of the cascade. Ideally a unique identifier would exist within the TB program, because surveillance data would then allow for identification of patients who never reached the next step of the cascade, potentially prompting active outreach efforts to re-engage them in care. This would also yield more accurate estimates as one could ensure that patients quantified at each step of the cascade represented a subset of the same patients quantified in the preceding step.

To address a point from Reviewer #1, we offered clarification around this point in the revised methods section, “*Furthermore, unique patient identifiers are not available within Zambia’s NTP and thus this analysis does not present a cohort of individuals that were tracked through each step of the TB care cascade; while we assumed for the purposes of this analysis that the same patients were being characterized at each step of the cascade, one cannot exclude the possibility that different individuals are being captured at different steps of the care cascade.”*

14.Line 4 – I suggest removing ‘among individuals’

Response: Thank you. This change has been made.

15.Line 46 – add ‘coma’ between ‘Zambia’ and ‘have’

Response: Thank you. This change has been made.

16.Line 47 – suggest citing the latest Global TB report, if possible.

Response: Thank you. This change has been made and corresponding statistics have been updated

17.Line 53-55: Review grammar for sentence construction and flow.

Response: Thank you. The sentence has been revised as follows, *"The HIV "cascade of care" is a public health model that outlines the several key engagement steps required for PLHIV to ultimately achieve an undetectable viral load."*

18.Line 60-62: Review the construction of the sentence

Response: Thank you. The sentence has been revised as follows, *"However, to-date, only three high burden TB countries - South Africa, India, and Madagascar - have undertaken and published national-level TB care cascade analyses."*

19.Line 66 – Consider not making rifampicin-susceptibility a single word; same for HIV-status

Response: Thank you. These changes were made.

20.Line 73 – Consider removing 'Living in its Provinces'

Response: Thank you. This change has been made.

21.Line 7 - There is a word missing

Response: Thank you. We have made this correction at line 74.

22.Line 81 - Indicate the basis of your estimation that a small number of patients are detected and managed in the private sector;

Response: Thank you for seeking clarification on this important assumption. We have revised the manuscript as follows, *"While exact estimates are not available, likely <1% of all TB cases are detected and managed within Zambia's private sector and the large majority are reported to Zambia's National TB Program (NTP) – this assumption is informed by a data quality audit conducted in 2019 (unpublished)."*

23.Line 88 – Consider changing 'multi-drug resistant' to become 'multidrug-resistant'

Response: Thank you. This change has been made.

24.Could the authors elaborate why information from only 4 provinces was used to draw national programmatic data?

Response: We thank the reviewer for this query. In response to Reviewer 2's query and that of Reviewer 1, we have added further details to the methods section of the revised manuscript. In brief, these 4 provinces were provided technical support by the Center for Infectious Disease Research in Zambia and therefore, well-characterized, patient-level, TB treatment data was available. However, these provinces account for nearly 60% of all TB case notifications and are felt to be representative of the country as a whole. The following additional details have been

added*, "Patient-level data was only available from 4 provinces of 10 provinces; however, they account for nearly 60% of Zambia's national TB notifications and the range of socioeconomic characteristics of individuals as well as their access to healthcare services are representative of the other 6 provinces. Unfortunately, robust data from 2018 to inform these estimates were unavailable – thus, we utilized 2017 data because it was well-characterized and temporally close to the year for which we sought to characterize the TB care cascade."*

25. Line 133: To improve intelligibility, please consider inserting coma just before 'and smear'

Response: Thank for this suggestion. We agree that this sentence was not as clear as it could have been and have revised it as follows, *"Diagnostic sensitivity estimates of Xpert and smear microscopy for the detection of TB stratified according to HIV status, as well as Xpert, molecular line probe assays and liquid culture for rifampicin-resistance were informed by previously published systematic reviews and meta-analyses."*

26. Line 184 - Consider elaborating "Xpert tests sent each year" - to where?

Response: Thank you for this suggestion. To improve readability of this section, we have revised the wording as follows, *"Using facility-level aggregated laboratory data, we plotted (a) the total number of sputum Xpert tests undertaken each year against the total number of pulmonary TB cases diagnosed each year, including the proportion that was microbiologically confirmed as well as (b) the total number of Xpert tests undertaken (on any specimen) each year against the total number of RR-TB cases diagnosed and notified each year."*

27. Line 219: Sentence starting 'The majority…' is missing a word.

Response: Thank you. This change has been made.

28.Line 227-229 – Consider skipping part after 'HIV,"

Response: Thank you. This change has been made.

29.Line 257: If 'treatment outcomes' is referring to 'Treatment completion rates', perhaps authors could consider phrasing consistently

Response: Thank you. This change has been made.

30. Line 297 – 'This has been…"; please specify 'what' has been

Response: Thank you for pointing this out. This has been revised as follows, *"Furthermore, once patients do access healthcare services, their TB illness may be missed – this has been shown to be a common problem in recent standardized patient studies conducted in Kenya, India, and China."*

31. Line 299 - check gramma

Response: *Thank you. This has been revised as follows, "In the last Zambian national TB prevalence survey from 2013/2014, only 60% of previously undiagnosed TB cases were symptomatic, of which 50% had sought care for their illness at a health facility."*

32. Line 350 – Consider correcting gramma

Response: Thank you. This has been revised as follows, *"PLHIV accounted for more than 60% of TB cases and Zambia and were more likely to be lost at several steps of the cascade compared to HIV-negative individuals."*

33. Line 355: Consider rephrasing to, especially regarding this part: "presenting to and in-care"

Response: Thank you. This has been revised as follow, *"Non-specific/mild symptoms may delay care-seeking among PLHIV, and without systematic TB screening of PLHIV presenting to care, diagnosis may be even further delayed."*

34. define LFTU at first use

Response: Thank you. This change has been made.

35. The outcomes figure shows overall improvement, but a downward trend for 3 of the 5 measures. Is this something the authors may consider commenting on even in passing; if there is something to read into it?

Response: We thank the reviewer for this comment. Indeed, two of the patient groups demonstrated small declines in the proportion completing treatment. We have highlighted this in the revised results section. Furthermore, to add improved clarity, we have revised figure 3 to include corresponding 95% confidence intervals.

VERSION 2 – REVIEW

REVIEWER	Loutet, Miranda Public Health England
REVIEW RETURNED	25-Mar-2021

GENERAL COMMENTS	Overall comments: This paper now reads extremely well and is much clearer. Thank you for addressing each of my comments in great detail and making the necessary changes. I have only added a few more comments to further clarify the methods and aid with ease of reading. - The acronym PTLTFU is very hard to remember and is not common in the literature to I would suggest writing out "LTFU prior to initiation of TB treatment" each time. Abstract - Be clearer about the definition of this indicator "among those who did not complete the care cascade" because the readers have not been introduced to definitions yet. Suggest changing to make it clear that you mean they did not complete treatment. Background

	- Line 45: incidence should just be written as “333 per 100,000” - Would also be interesting to add a line of how that compares to other countries in SSA or global TB estimates. Methods - Line 77-78: why is this not published? Not robust to cite unpublished literature, so suggest adding some citable evidence e.g., government reports. - Line 94: remove “all” from this sentence. Your care cascade shows that not all patients diagnosed with TB are initiated on TB therapy. - Line 153-154: the acronym for DS-TB is defined twice here. Also, why is HIV status only known among patients with DS-TB? This is also a very important outcome among patients with RR-TB. If this is an issue of availability of data, then that should be stated here in the methods. - Line 154: be consistent using the abbreviation for rifampicin resistant TB i.e., RR-TB throughout the entire manuscript. - Line 161: add reference to the WHO recommendations Results - Line 226 - 227: be consistent with using “range:” because on line 227 you have used “range,” - Line 283: specify “RR-TB cases” - Line 284: here again be consistent with the use of the abbreviation for RR-TB - Table 2: I appreciate seeing the two proportions in the same table but it is confusing they are both under the same title. Please include a short description of the % in the title heading of the table (e.g., %total burden, %cascade step) and also provide the description in a footnote of the table as is already done. However, use a new symbol for this footnote, as it is different from the “cases range” footnote. Discussion: - line 351: missing a period after Zambia - line 410: the authors cannot generalize results to all TB cases in Zambia because estimates were only among DS-TB patients. Also suggest saying “60%” rather than “more than…” because it was only 61%. Therefore, the authors need to be specific here and suggest rephrasing to “PLHIV accounted for 60% of DS-TB cases in Zambia”. The authors could then add a sentence here that it is most likely higher than 60% because estimates were not able to include RR-TB patients (and explain why you could not include estimates among RR-TB patients). - Line 434: here again use abbreviation for RR-TB

REVIEWER	Chikovore, Jeremiah Human Sciences Research Council
REVIEW RETURNED	18-Apr-2021

GENERAL COMMENTS	Congratulations to the authors who have addressed concerns and queries from my earlier review. May I request that they re—look or correct a few issues, as specified below.

	1. Line 16-18: “Among those who did not complete the care cascade, 73.1% were lost prior to accessing diagnostic services, 8.1% prior to diagnosis, 9.4% prior to initiating treatment and 9.4% prior to treatment completion”. – Can authors include the absolute numbers too. 2. Line 296-297: “The proportion of HIV-positive patients completing TB therapy remained relatively unchanged from 2015 to 2018 (87.3% vs. 88.4%, 358 p=0.001)” Can the authors verify this is written as intended, paying attention to the stated p-value? 3. Line 375-377 - needs correction. “A recent qualitative study among TB patients and health care workers (HCW) in India provided further understanding of factors that may contributed to PTLTFU [48][48].” 4. Please pay attention to distortions to references in the list possibly from using a reference managing program. Also be consistent about how an author is named across source documents originating from the same author.

VERSION 2 – AUTHOR RESPONSE

REVIEWER #1

3. This paper now reads extremely well and is much clearer. Thank you for addressing each of my comments in great detail and making the necessary changes. I have only added a few more comments to further clarify the methods and aid with ease of reading.

Response: Thank you very much. We are glad that the reviewer believes that we have adequately addressed their prior points and that the paper has been strengthened. We also thank them for their additional review and helpful comments.

4. The acronym PTLTFU is very hard to remember and is not common in the literature to I would suggest writing out “LTFU prior to initiation of TB treatment” each time.

Response: Thank you. This has been revised throughout the manuscript according to the reviewer’s suggestion.

5. Abstract: Be clearer about the definition of this indicator “among those who did not complete the care cascade” because the readers have not been introduced to definitions yet. Suggest changing to make it clear that you mean they did not complete treatment.

Response: Thank you. To further improve clarity, we have further revised this sentence as follows, “Among all TB patients lost at any step along the care cascade (n=39,795), 29,108 (73.1%) were lost prior to accessing diagnostic services, 3,211 (8.1%) prior to diagnosis, 3,745 (9.4%) prior to initiating treatment and 3,731 (9.4%) prior to treatment completion.”

6.Background: Line 45: incidence should just be written as “333 per 100,000.”

Response: Thank you. We have revised this sentence accordingly. “In 2019, there were approximately 59,000 new individuals with active TB disease in Zambia (incidence rate of 333 per 100,000 per year).”

7. Would also be interesting to add a line of how that compares to other countries in SSA or global TB estimates.

Response: Thank you. We have added the following sentence, "Despite substantial declines in TB incidence over the last decade, Zambia still has the seventh highest TB incidence in sub-Saharan Africa and remains one of 30 WHO high TB burden priority countries."

8. Methods: Line 77-78: why is this not published? Not robust to cite unpublished literature, so suggest adding some citable evidence e.g., government reports.

Response: Thank you. We have now added a reference from a government report as recommended by the reviewer.

9. Line 94: remove "all" from this sentence. Your care cascade shows that not all patients diagnosed with TB are initiated on TB therapy.

Response: Thank you. This has been revised accordingly.

10.Line 153-154: the acronym for DS-TB is defined twice here. Also, why is HIV status only known among patients with DS-TB? This is also a very important outcome among patients with RR-TB. If this is an issue of availability of data, then that should be stated here in the methods.

Response: Thank you. We have removed the duplicate mention to DS-TB. With respect to characterizing the care cascade among RR-TB disaggregated according to HIV status, we agree that this would be informative to characterize. Unfortunately, due to the way in which available data sources are captured and reported, we are not able to undertake this analysis. To make this clear, we have added the following sentence to the manuscript as advised by the reviewer, "There was insufficient data available to characterize the RR-TB care cascade disaggregated according to HIV status."

11. Line 154: be consistent using the abbreviation for rifampicin resistant TB i.e., RR-TB throughout the entire manuscript.

Response: Thank you. This has been revised accordingly.

12. Line 161: add reference to the WHO recommendations

Response: Thank you. The appropriate reference has been added.

13. Results: Line 226 - 227: be consistent with using "range:" because on line 227 you have used "range,"

Response: Thank you. We have carefully reviewed the results section to ensure that this is applied consistently throughout.

14. Line 283: specify "RR-TB cases"

Response: Thank you. This has been revised accordingly.

Line 284: here again be consistent with the use of the abbreviation for RR-TB

Response: Thank you. This has been revised accordingly.

15. Table 2: I appreciate seeing the two proportions in the same table but it is confusing they are both under the same title. Please include a short description of the % in the title heading of the table (e.g., %total burden, %cascade step) and also provide the description in a footnote of the table as is already done. However, use a new symbol for this footnote, as it is different from the "cases range" footnote.

Response: Thank you for these helpful suggestions. We agree that this provides greater clarity and these recommendations have been incorporated into the revised Table 2.

16. Discussion: line 351: missing a period after Zambia

Response: Thank you. This has been revised accordingly.

17. line 410: the authors cannot generalize results to all TB cases in Zambia because estimates were only among DS-TB patients. Also suggest saying "60%" rather than "more than…" because it was only 61%. Therefore, the authors need to be specific here and suggest rephrasing to "PLHIV accounted for 60% of DS-TB cases in Zambia". The authors could then add a sentence here that it is most likely higher than 60% because estimates were not able to include RR-TB patients (and explain why you could not include estimates among RR-TB patients).

Response: Thank you. We have revised this sentence as follows, "PLHIV accounted for 60% of DS-TB cases in Zambia and were more likely to be lost at several steps of the cascade compared to HIV-negative individuals." While we agree that after accounting for RR-TB, estimates may be slightly higher, because RR-TB accounts for only 2-3% of all TB cases, it would be unlikely to meaningfully change overall estimates, and without further data, the directional impact on estimates would be subject to speculation. With regard to the inability to include estimates of RR-TB according to HIV status, we have revised the methods section in response to point 10 above.

18. Line 434: here again use abbreviation for RR-TB

Response: Thank you. This has been revised accordingly.

REVIEWER #2

19. Congratulations to the authors who have addressed concerns and queries from my earlier review. May I request that they re—look or correct a few issues, as specified below.

Response: Thank you very much. We are glad that the reviewer believes that we have adequately addressed their prior points and that the paper has been strengthened. We also thank them for their additional review and helpful comments.

20. Line 16-18: "Among those who did not complete the care cascade, 73.1% were lost prior to accessing diagnostic services, 8.1% prior to diagnosis, 9.4% prior to initiating treatment and 9.4% prior to treatment completion". – Can authors include the absolute numbers too.

Response: Thank you. We have changed this sentence accordingly. "Among all TB patients lost at any step along the care cascade (n=39,795), 29,108 (73.1%) were lost prior to accessing diagnostic services, 3,211 (8.1%) prior to diagnosis, 3,745 (9.4%) prior to initiating treatment and 3,731 (9.4%) prior to treatment completion."

21. Line 296-297: "The proportion of HIV-positive patients completing TB therapy remained relatively unchanged from 2015 to 2018 (87.3% vs. 88.4%, 358 p=0.001)" Can the authors verify this is written as intended, paying attention to the stated p-value?

Response: Thank you for seeking clarification on this point as we agree that it is not as clear as it could be. The p-value is correct. The values differ statistically because of the large denominator, but they do not meaningfully differ. To make this clearer, we have revised this sentence as follows, "The proportion of HIV-positive patients completing TB therapy did not meaningfully change from 2015 to 2018 (87.3% [95%CI: 86.9-87.7] vs. 88.4% [95%CI: 88.0-88.9]; p=0.001)."

22. Line 375-377 - needs correction. "A recent qualitative study among TB patients and health care workers (HCW) in India provided further understanding of factors that may contributed to PTLTFU [48][48]."

Response: Thank you. We have revised the sentence accordingly. "A recent qualitative study among TB patients and health care workers (HCW) in India provided further understanding of the factors that may contribute to LTFU prior to the initiation of TB therapy [49]."

23. Please pay attention to distortions to references in the list possibly from using a reference managing program. Also be consistent about how an author is named across source documents originating from the same author.

Response: Thank you. We have gone through the reference list carefully and revised references wherever appropriate.

VERSION 3 – REVIEW

REVIEWER	Loutet, Miranda Public Health England
REVIEW RETURNED	24-Jun-2021

GENERAL COMMENTS	The edits to the paper have greatly improved the clarity of the paper. It is now easy to follow through the methods, results and discussion. There are just a few minor comments outlined below. Specific comments: Methods – line 113-117: I am not convinced that this section is necessary in this scientific paper. Is this a requirement for BMJ open? If not, then I think you can remove it as it only adds more bulk to the paper. Discussion – line 488: why is the word "core" used here to describe incidence? Unless there is a particular reason, then suggest removing that word as it is not an epidemiological term.

REVIEWER	Chikovore, Jeremiah Human Sciences Research Council
REVIEW RETURNED	07-Jun-2021

GENERAL COMMENTS	The authors have addressed my outstanding queries to my satisfaction. A minor comment that needs addressing is the presentation of the Methods section. Line 68-72 seems to present Methods. The sub-title 'Methods' is not distinguishable as higher level to 'Setting' and other subtitles that ordinarily should fall under 'Methods'. This implies Lines 68-72 is a standalone section, and maybe the entire methods section.

VERSION 3 – AUTHOR RESPONSE

REVIEWER #1

3. The edits to the paper have greatly improved the clarity of the paper. It is now easy to follow through the methods, results and discussion. There are just a few minor comments outlined below.

Response: Thank you very much. We are glad that the reviewer believes that we have adequately addressed their prior points and that the paper has been strengthened. We also thank them for their additional review and helpful comments.

4. Methods – line 113-117: I am not convinced that this section is necessary in this scientific paper. Is this a requirement for BMJ open? If not, then I think you can remove it as it only adds more bulk to the paper.

Response: Thank you for this suggestion. It is our understanding that this section is a requirement of BMJ Open. However, as suggested by the Editor, we have moved this section out of the main text body within the methods section and moved it to the end of the manuscript.

5. Discussion – line 488: why is the word “core” used here to describe incidence? Unless there is a particular reason, then suggest removing that word as it is not an epidemiological term.

Response: We agree with this suggestion. We have removed the word “core” as suggested by the reviewer.

REVIEWER #2

6. The authors have addressed my outstanding queries to my satisfaction.

Response: Thank you very much. We are glad that the reviewer believes that we have adequately addressed their prior points. We also thank them for their additional review and helpful comments.

7. A minor comment that needs addressing is the presentation of the Methods section. Line 68-72 seems to present Methods. The sub-title 'Methods' is not distinguishable as higher level to 'Setting' and other subtitles that ordinarily should fall under 'Methods'. This implies Lines 68-72 is a standalone section, and maybe the entire methods section.

Response: We thank the reviewer for this suggestion, and we agree that this may be unclear. To improve clarity, we have made the following two changes. First, we added a sub-header, "Study design," for the first paragraph of the methods section. Second, all sub-headers within the manuscript have been changed from bolded to underlined text. We hope that this now makes it clear that "Methods" is a higher level than the "Study design" sub-section, and that headers and sub-headers are clearer throughout the entire manuscript.